Volume 209, number 1 FEBS 4250 December 1986

The similarities and differences in structures between kringle 1 of prothrombin and kringle 4 of plasminogen

R.J.P. Williams, P. Esnouf, M. Lawrence and S.A. Cederholm-Williams

Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3 QR, England

Received 16 October 1986

Abstract not received

Kringle structure *NMB* *Prothrombin* *Plasminogen* *β-sheet* *Lysine site* *Protein mobility*

1. INTRODUCTION

Two structures of protein kringles (fig.1) have been published. The outline structure of kringle 4 (plas) of plasminogen was published by us using detailed NMR analysis [1]. This structure has been confirmed in almost all respects, see below, by much further NMR work [2]. An outline structure of kringle 1 (pro) of prothrombin has been published recently using X-ray diffraction at about the same level of detail as we now have for the kringle 4 [3]. A very satisfactory feature of these two studies, which were done totally independently (neither group knew of the work of the other), is the level of agreement. Fig.2 reproduces the main chain fold from kringle 1 (pro) since that structure is derived more largely from chain following. The NMR study had to be based heavily on cross-chain interactions since sequential NMR methods [4] were only partially viable [1]. [The difficulties in the NMR study become more relevant later.] A comparison of the two is most easily given therefore by reference to the X-ray fold features of fig.2.

2. A COMPARISON BETWEEN KRINGLE 4 (PLAS) AND KRINGLE 1 (PRO)

The two structures illustrated in both papers have obvious similarity of overall dimensions and both have a small protruding section where the N-terminus (residue 1) and the C-terminus (residue 79) come together at an -S-S- bridge outside the main protein envelope (fig.2). [We use kringle 4 (plas) numberings from the first cysteine, fig.1.] The interaction between the chains continues as seen by cross-connections from Tyr-2 to Leu-76 in the NMR spectrum. Both structures have a small region of antiparallel β-pleated sheet across from residue Phe-63 which lies in a β-strand from 61 to 65 but there is no clear α-helix in either structure. Here there are some differences already, however, in that the sheet is differently composed in the two structures. It is apparently triple stranded in kringle 4 (plas) [1] but only of two strands in kringle 1 (pro) [3]. The cross-strand residues involved are also different. [We will return to this difference later since the β-sheet is not all that stable or well-defined as judged by NMR data.] The positions of turns (β-turns largely at prolines) in the stretches 22–50 (second long arm), 51–62 (first short arm) and 63–74 (second short arm) are very similar indeed. There are also turns in the first long arm at around residue Tyr-2 and 15–18 but the last group of residues is not so similarly placed in the two kringles. The middle of this arm is ill-defined in the NMR study, see below. The same cross connections between remote parts of the chain are described in both publications such that there is a relationship between the looped sequence

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

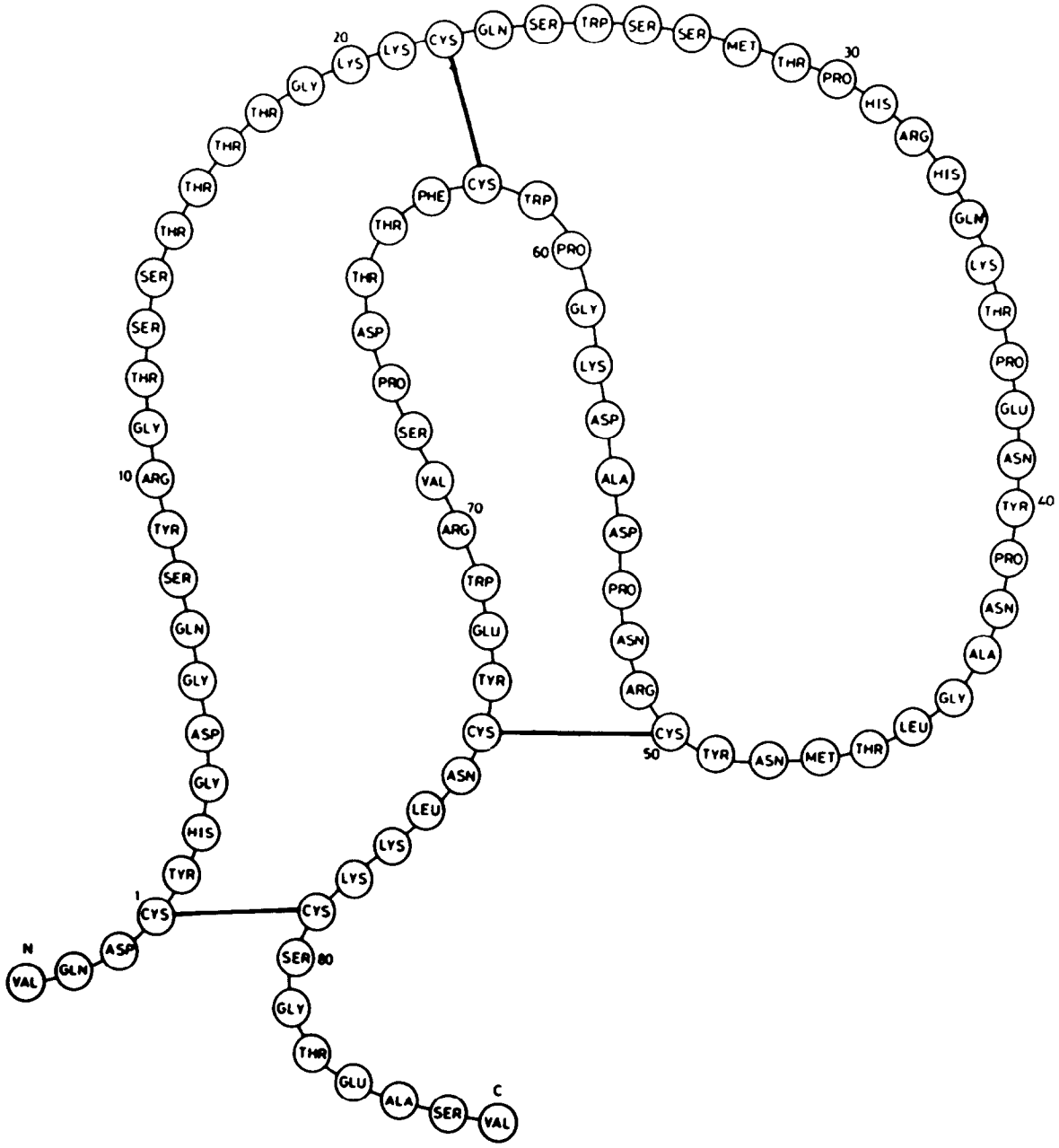

Fig.1. The kringle domain as determined from sequence studies (see [1] for early references). The amino acids shown are those of plasminogen kringle 4. Table 1 shows regions of the sequences of several other kringles stressing common features other than the positions of the three -S-S- bridges.

from 23 to 35 which brings it close to, and interacting with, a similar loop from 62 to 74. There is possibly a strong H-bond from His-31 (wrongly listed in [1] as His-C – it is His-A of previous NMR papers [5]) to Asp-66 which would account for the very low apparent pK_a of 4 of His-31. These are the two loops seen on the right-hand side of the structure of kringle 1 (fig.2) [3] and were similarly

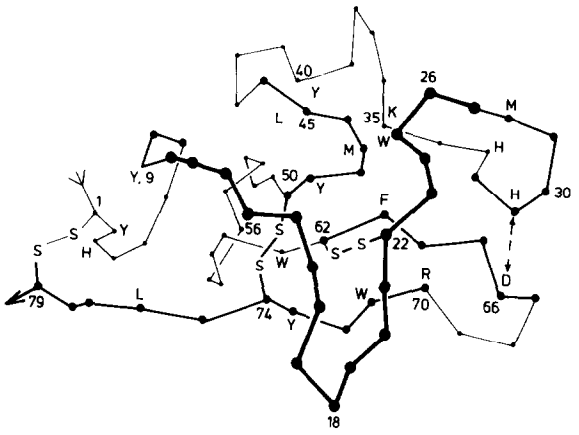

Fig.2. The kringle fold of prothrombin kringle 1 with the sequence of plasminogen kringle 4 imposed on it. The diagram is adapted from [3]. Features of note are the two inner shorter runs of fig.1 which fold together so that much of the top of fig.1 comes close to the bottom, e.g. in one cluster Trp-25 and Leu-45, and in another Trp-61 and Phe-63 come together with Tyr-73 and Trp-71 (antiparallel β-sheet in kringle 4 (plas) with all four aromatic side-chains on the same side giving a hydrophobic back to the receptor site). The back and bottom of the diagram, involving the sequence 51–60 in a pair of turns, carries the carboxylic acid pair, Asp-54 and -56 ($pK_a = 5.0$), which give the cationic site while the Arg-70 controlled by the above four aromatic residues gives the anionic site for ligands $H_3N^+ \cdot X \cdot CO_2^-$ where X is a hydrophobic chain of some 6–8 carbon length and rests on the above hydrophobic residues. Overall the structure is a rough sphere with (a) a protruding arm, 1–8 and 75–79, terminating in an -S-S- bridge and leading to the connecting mobile sequences between kringles 3 and 4 and between kringles 4 and 5, and (b) a concave indentation which is the receptor site. The convex surfaces are donors to other receptors of the complex organisation in blood clot removal. All other kringles are similar but the differences between kringle 4 (plas) and kringle 1 (pro) suggest that some care is necessary in making comparisons.

disposed in our model [1]. The position of the chain from 35 to 50 (apart from the agreement on folding which turns out to be very close to our new NMR results, which were being written up [2] when the X-ray structure appeared) is virtually the same in both structures. Leu-45 in both, and indeed in at least four kringle structures examined by NMR [1,4], sees residue 40 (Tyr) and residue 40 sees residue 35 (Lys). Leu-45 (constant in all kringles) is also close to Trp-25 (constant in all kringles). In both structures there is an aromatic cluster of three residues Phe(Tyr)-63, Trp-61 and Trp-25, all of which see one another, i.e. they have protons less than 4–5 Å from one another as judged by NMR studies. A difference does appear however in that the orientation of Trp-61 which points more with its C-2 in the direction of two aspartates 54 and 56 in the first short chain in kringle 4 (plas). The orientation is apparently in almost the opposite sense in kringle 1 (pro) [2]. These aromatic groups are not very exposed. The relationship of other aromatic groups and the exposure of them is also identical with the possible exception of Tyr-9 which is difficult to place by NMR for reasons to be described below. It is certainly buried. Further examples of agreement are that the exposed Tyr-49 sees Thr-46 and Thr-12, and hidden Tyr-73 protons show NOE cross peaks to Trp-61 showing it to lie close to Trp-61. In kringle 4 (plas) we can also place the third tryptophan, Trp-71 (absent in kringle 1), interacting through its α-CH with the α of Phe-63 in the β-sheet and through its C-2 to stabilise Arg-70. Substrate binding of kringle 4 is associated with the stabilisation of Arg-70. Both structures leave a binding concave pocket across from one short segment around Arg-70 to another around Asp-54 and Asp-56. These are the binding groups for recognition of the electrostatically held lysine-like molecules in adducts of kringle 4 (plas). [Kringle 4 (plas) is a lysine binding domain.] We discuss this binding domain further below. The NMR spectra of other kringles have similar characteristics. We now see that the -S-S- links, at least one β-(sheet) strand, the hydrophobic clusters and many β-bends are common and in fact the NMR data show that they are common to all kringles so far examined [1]. This is a somewhat expected conclusion even from sequences (table 1).

3. THE DIFFERENCES BETWEEN THE TWO STRUCTURES

Apart from the small differences mentioned above which could arise from errors of very different kinds in the two methods there are four considerable differences between the two kringle structures but before describing them we notice that the description of kringle 4 (plas) and indeed other kringles by NMR meets several problems.

Table 1

Sequences of some constant regions of kringles

	48		50										60	
BF1	N	F	C	R	N	P	D	G	S	I	T	G	P	W
HF1	–	–	–	–	–	–	–	S	S	I	T	–	–	–
BF2	–	–	–	–	–	–	–	G	D	E	E	–	A	–
HF2	–	–	–	–	–	–	–	G	D	E	E	–	V	–
PK1	–	Y	–	–	–	–	–	N	D	P	Q	–	–	–
PK2	–	Y	–	–	–	–	–	R	E	L	R	*	–	–
PK3	–	Y	–	–	–	–	–	G	K	R	A	*	–	–
PK4	–	Y	–	–	–	–	–	A	D	K	G	*	–	–
PK5	–	Y	–	–	–	–	–	G	D	V	G	–	–	–
PA1	–	Y	–	–	–	–	–	R	D	S	K	*	–	–
PA2	–	Y	–	–	–	–	–	G	D	A	K	*	–	–
PA3	–	Y	–	–	–	–	–	N	R	G	R	*	–	–

	62								70				74	
BF1	C	Y	T	T	S	P	T	L	R	R	E	E	C	
HF1	–	–	–	–	D	–	–	A	–	–	Q	E	–	
BF2	–	–	V	A	D	Q	P	G	D	F	–	Y	–	
HF2	–	–	V	A	G	K	P	G	D	F	G	Y	–	
PK1	–	–	–	–	D	–	E	K	R	Y	D	Y	–	
PK2	–	F	–	–	D	–	N	K	R	W	–	L	–	
PK3	–	H	–	–	N	S	Q	V	R	W	–	Y	–	
PK4	–	F	–	–	D	–	S	V	R	W	–	Y	–	
PK5	–	–	–	–	N	–	R	K	L	Y	D	Y	–	
PA1	–	–	V	F	K	A	G	K	Y	S	–	F	–	
PA2	–	H	V	L	K	N	R	L	T	W	–	Y	–	
PA3	–	–	V	V	G	L	K	P	L	V	Q	E	–	

*, absent residue; invariants apart from those shown above: C, at positions 1, 22, (50), (62), (74) and 79; G, at 5, 11, 19; Y, at position 9; W, at position 25 and 61; L, at position 45. The kringle symbols are for bovine and human prothrombin (BF and HF), plasminogen (PK) and plasminogen activator (PA). Particularly observe the changes at 71 and 73 and the changed β-sheet structure between BF1(3) and PK4(1). This could be a rare case where some elements of secondary structure are not preserved in very homologous proteins

(I) There are pK_a values close to pH 6 due to Asp-54/56 and histidine 3 and 33 ionisations which cause some conformational changes and the fold is disrupted at low pH where the protein precipitates. Trp-61 resonances for example are greatly affected by pH change even going from pH 6 to pH 7. At all working pH values more than one conformation contributes to the NMR spectrum. We cannot work in the best conditions to avoid NH exchange and acid-base equilibria and yet get a pH 7 structure. These problems are inherent in NMR methods when studying a dynamic structure and are not so obvious in crystallographic studies due to crystal constraints. The pH in the crystals of kringle 1 (pro) is not reported [3].

(II) There are broad lines in some parts of the structure especially Tyr-9 is very difficult to follow as if it exists in more than one state, perhaps intermediate rate of flipping. Broad lines are common to Phe-63 and some (not all) threonines which may well arise from instability of the β-strand region. The uncertainties in the crystallographic study appear to be much smaller but we must wait for the publication of electron density maps and coordinates.

The four major differences between the published structures are as follows.

(i) The apparent β-sheet region has three strands by NMR as published [1] but only two by X-ray [3] and the sheet is very differently described. The X-ray structure puts it between sequence regions 48–53 and 59–64 (kringle 4 (plas) numberings) while the NMR study [1] placed it between the sequence 12–20 and 63–64 and an unknown region but now known to be 71–73. The NMR β-sheet structure is now known to be due to the antiparallel run of 61–65 and 69–73 while there is a strand crossing of 15–18 over 62–65 on the opposite side of 61–65. If both structures are correct the 'β-sheet' region is variable in kringles although both are based on the β-strand around aromatic residue 63. There is evidence from NMR for this variability even in kringle 4 (plas) in that some of the cross-contacts just described are sensitive to pH. The core of the sheet based on residue Phe-63 and Trp-61 is the same in all structures. We know this for kringle 2 (pro) and kringle 1 (plas) too. The H-bond network of the crystal structure is needed for a proper comparison.

(ii) The distance between the two -S-S- bridges, Cys-50–Cys-74 and Cys-22–Cys-62, is very small in the prothrombin kringle 1 (pro) crystal structure – only 4 Å at the closest [3]. In our published model the distance is 7 ± 1 Å in kringle 4 (plas) [1] but small changes in bond angles could reduce this possibly to 6 ± 1 Å (The NMR definition of this relationship is not fully explored). These two -S-S- bridges are in the centre of the fold in both structures but in the published NMR study they were not placed relative to one another because of assignment problems of cysteines but were placed by the fold characteristics which limited their position grossly in space. As with the difference (i) above more detailed study is needed especially here in the NMR study where dihedral angles are being carefully analysed. The absence of much well-defined secondary structure makes the NMR study difficult.

(iii) The third difference is also illuminating. It concerns the first long chain from 1 to 22. There is disagreement as to the disposition of this chain in the two kringles from about residue 12 (Thr) to 22 (Cys) itself. Residue 9 (Tyr) is not easily placed in kringle 4 (plas) by NMR methods but there follows a sequence from residue 12 to 18 which contains five threonines and is easily studied. The lack of simple secondary structure here, although there is a turn at residues 17–19 amongst the four threonines, as in kringle 1 (pro), prevented us from defining this region closely in our earlier paper on kringle 4 (plas) but new data show that the chain at residue 12 (threonine) passes close to Tyr-49 as in kringle 1 (pro) but it then interacts with the residues of the β-strand 62–65 such that Thr 15 or 16 has a (β-sheet type) NH hydrogen bonding to the β-strand 63–65 region. The close interaction of three threonines all close to Phe-63 [1] is indisputable in kringle 4 (plas) but is absent apparently in kringle 1 (pro) of prothrombin although both have a threonine at this position. Thus this description is quite different from the chain disposition in kringle 1 [3] where residues 10–22 approach the -S-S- link at 22 from what appears to be the opposite side of Phe-63 via a wide turn carrying a carbohydrate at position 12, a straight stretch to position 17 and a subsequent turn at residue 18 and 19. This is the chain disposition in fig.2. The connections to the β-strand 62–65 are clearly different. The position of this long chain 1–22 remains on the surface (see fig.1). This change may make an adjustment of the -S-S- bridge (22–62) which could explain (ii) above. The further analysis of these features could help to explain the selectivity of action of the kringles, see below.

We shall publish all assignments and structural data shortly but the differences described above cannot be regarded as reconcilable.

(iv) There is one other (amusing) difference. Our published structure [1] is clearly almost the mirror image of kringle 1 (pro) [3]. The chain around residue 55 is at the bottom and at the front in our picture (fig.7 of [1]). [Unfortunately the reproduction is not clear.] It is an intrinsic problem of NMR that optical isomers are not resolvable unless one starts from some known chiralities. We did not do this nor could we have done so at the early stages using a distance-geometry approach without much use of sequential data. No real problems arise from this mistaken handedness at this stage since we are not concerned with the binding of optically active ligands.

4. FURTHER NMR KNOWLEDGE

The original model we published was greatly helped by the use of lanthanide shift and relaxation probes [1]. The cavity for ligands was given a relatively well-defined geometry based on the binding of Ln(III) ions to Asp 54 and Asp 56 which is known to occur without any real conformation changes. We have now observed the binding of amino methylene benzoic acid, an antithrombitic drug, across to Arg-70 from those two aspartates. The ring of this model for lysine, the natural ligand, acts as a new shift probe. It interacts as expected with Trp-61, Trp-71 (less), and Phe-63 at the back of the hydrophobic pocket but with very few aliphatic residues. The fitting of the ligand is exactly as predicted [1] and the ligand resonances themselves are also shifted in an expected manner. [We have been informed that Dr L.R. Brown (Canberra) and Dr M. Llinas (Illinois) have made similar observations.] However it is also observed that the apparent β-sheet region is greatly tightened since α-NH protons of the region do not exchange now over a period of days.

5. INFERENCES CONCERNING KRINGLE STRUCTURES

There is clearly a fixed pocket formed from the two short strands 50–62, 62–74 of both, probably all, kringles together with at least both end-parts of the second long strand 22–40 which act as hydrophobic supports, at Leu-45 and Trp-25. This is the binding pocket of kringle 4 (plas). This concave surface, back of fig.2, backed by hydrophobic residues must be common to all kringles, see table 1. Residues 25, (27), 45, 49–55, 59–64 and all six cysteines are conserved. The ligand specificity is then built into variations of short parts of the binding regions 56–58 and 65–71. Little variation in fold appears to be possible here and there is restricted mobility of many side chains due to internal packing. This opens up many possibilities for drug design and bio-engineering since kringles are essential for blood clotting and clot removal. Note the variations in table 1 against fig.2. It must be a relatively easy matter to design potential drugs, binding agents, since the kringles have the same hydrophobic concave surface and the separation between the hydrophilic binding centres Arg-70 to Asp-54 and -56 is about equivalent to a 6–8 atom, largely carbon, chain as in lysine $H_3N^+ \cdot CH_2 \cdot CH_2 \cdot CH_2 \cdot CH_2 \cdot CO_2^-$. Importantly the end hydrophilic groups of the drugs could be varied to match the changes around sequence positions 70 and 55. These points will be explored elsewhere but notice the striking differences from the antibody binding domains. In kringles the constant domain is the basis of the binding pocket and only its lips are variable. More variable are the surfaces away from this pocket especially of the two long chains.

We turn to the convex surfaces and the more exposed regions. The protruding lobe Cys-1 and Cys-79 is artificial since it is connected to pieces of sequence removed by proteolysis. Some parts of this region have been discussed in both earlier publications [1,4]. There is undoubtedly mobility just beyond these regions [1,2] and the extra-kringle sequences are readily attacked by proteases in the breakdown of the pro-proteins, prothrombin and plasminogen. We do not comment further on the rather variable sequences 2–8 and 75–78. Most interestingly we are left with two very variable regions from residue 10 to 22 (some structural differences) and from 28 to 44 (very similar folds). These are the two sequences which carry saccharides in kringle 1 (pro). They give another easy target for bio-engineering but note that the carbohydrate is disordered in the kringle 1 (pro) structure [2]. It will be interesting to see the ways in which protein/protein bindings in the complicated multiprotein units used in clot-forming, prothrombin, and clot-dissolving, plasminogen, owe their origins to these exposed surface regions which are so different. The contrast with the constant surface regions of the antibodies could not be greater.

REFERENCES

[1] Esnouf, P., Lawrence, M.P., Mabbutt, B.C., Patthy, L., Pluck, N.D. and Williams, R.J.P. (1985) Bull. Soc. Chim. Belg. 94, 883–896.
[2] Mabbutt, B. and Williams, R.J.P., to be published.
[3] Park, C.H. and Tulinsky, A. (1986) Biochemistry 25, 3977–3982.
[4] De Marco, A., Pluck, N.D., Banyai, L., Trexler, M., Laurssen, R.A., Patthy, L., Llinas, M. and Williams, R.J.P. (1985) Biochemistry 24, 748–754.
[5] Wider, G., Macura, S., Anil-Kumar, R., Ernst, R.R. and Wuthrich, K. (1984) J. Magn. Res. 56, 207–234.